Arch Virol (1996) 141: 197–204

Archives of Virology

Nucleotide sequence of the capsid protein gene of papaya leaf-distortion mosaic potyvirus*

Brief Report

T. Maoka[1], S. Kashiwazaki[2], S. Tsuda[3], T. Usugi[1], and H. Hibino[2]

[1] Okinawa Subtropical Station of the Japan International Research Center for Agricultural Sciences, Okinawa, [2] National Agriculture Research Center, Ibaraki, [3] Plant Biotechnology Institute, Ibaraki Agricultural Center, Japan

Accepted August 17, 1995

Summary. The DNA complementary to the 3′-terminal 1 404 nucleotides [excluding the poly(A) tail] of papaya leaf-distortion mosaic potyvirus (PLDMV) RNA was cloned and sequenced. The sequence starts within a long open reading frame (ORF) of 1 195 nucleotides and is followed by a 3′ non-coding region of 209 nucleotides. Capsid protein (CP) is encoded at the 3′ terminus of the ORF. The CP contains 293 residues and has a M_r of 33 277. The CP of PLDMV exhibits 49 to 59% sequence similarity at the amino acid level to the CPs of papaya ringspot potyvirus (PRSV) and other potyviruses. This result is consistent with the absence of a serological relationship between PLDMV and PRSV or other potyviruses. The results support the assignment of PLDMV as a distinct member of the genus *Potyvirus*.

*

Papaya (*Carica papaya* L.) is widely grown in the southern part of Japan. A viral disease of papaya was identified in 1954 on Okinawa Island and spread throughout the island during the 1960s [10]. Infected papaya plants show mosaic symptoms and distortion of leaves, as well as ringspots on fruits. These symptoms are similar to the symptoms caused by the virus known as papaya ringspot potyvirus papaya strain (PRSV-P) [10, 15]. PRSV-P is the major impediment to stable production of papaya fruits in many countries, including the U.S.A. and countries in South America, the Caribbean and

* Sequence data from this article has been deposited with the EMBL/GenBank/DDBJ databases under accession no. D50082.

Asia [15]. Papaya production has declined to a low level in many of these countries as a consequence of this virus. The papaya virus in Okinawa was first identified as PRSV-P from its similarity to previously characterized PRSV-P in terms of symptoms of infection of papaya plants, host range, relationship to aphid vectors and physical properties [10, 15]. Later, however, the virus was found to be distinct serologically from PRSV-P and the name "papaya leaf-distortion mosaic potyvirus (PLDMV)" was proposed for the virus [10].

PLDMV was considered to be the only virus that damages papaya plants in Japan until the occurrence of PRSV-P was discovered in 1992 [12]. However, PRSV-P was detected in only about 3% of 383 field samples tested by ELISA for both viruses [12].

The complete genomic sequence has been determined for the HA isolate of PRSV-P [25], and the capsid protein (CP) genes of some other isolates of PRSV have been sequenced [2, 3, 17, 21]. However, no sequence data is available for PLDMV RNA. Here we report the nucleotide sequence of the 3'-terminal region of the PLDMV RNA, which includes the CP gene, and we assess the taxonomic relatedness of PLDMV to PRSV and other potyviruses.

The P56 isolate of PLDMV [12] was propagated *in Cucumis metuliferus* by mechanical inoculation and purified as follows. Infected leaves were triturated in four times their weight of 0.5 M citrate buffer (pH 7.0) containing 0.1 M disodium hydrogen phosphate, and then the homogenate was squeezed through cotton cloth. The filtrate was clarified by adding carbon tetrachloride to 6%, stirring for 1 min at room temperature, and centrifugation at 6 000 × g for 15 min. Polyethylene glycol 6 000 (PEG), NaCl and Triton X100 were added to the supernatant to final concentrations of 7% (w/v), 0.1 M and 2% (w/v), respectively. The mixture was stirred for 30 min at 4 °C and centrifuged at 6 000 × g for 15 min. The pellet was resuspended in a volume equal to one-tenth the volume of the clarified supernatant of 0.1 M citrate buffer (pH 7.0) that contained 0.01 M disodium hydrogen phosphate (CD buffer) and centrifuged at 6 000 × g for 15 min. The resultant supernatant was centrifuged at 125 000 × g for 90 min. The pellet was resuspended in CD buffer and the mixture was centrifuged at 6 000 × g for 15 min. The resultant supernatant was layered on a 10–41% (w/v) linear density gradient of a cesium sulfate in CD buffer and centrifuged at 175 000 × g for 15 h in an RPS-40T rotor (Hitachi, Japan). The zone that contained the virus, located about one-third of the way up the centrifuge tube, was collected with a Pasteur pipette, diluted with CD buffer, and centrifuged at 238 000 × g for 90 min. The pellet was resuspended in 0.01 M citrate buffer (pH 7.0). The suspension of virus was used for our initial attempt to isolate viral RNA. However, electrophoresis of the RNA on the agarose gel resulted in smeared bands of DNA, which probably originated from host tissues and had, most likely, combined with PLDMV particles during the purification process. To remove the DNA, the suspension of virus was incubated with RNase-free DNase I (Boehringer Mannheim, Germany) at 1 500 unit/ml in 5 mM magnesium sulfate at 37 °C for 1 h. The DNase-treated suspension of virus was loaded on a cushion of 40% sucrose in

0.01 M citrate buffer (pH 7.0) and centrifuged at 128 000 × g for 60 min. The final pellet, containing virus, was suspended in small amount of 10 mM Tris-HCl (pH 8.0) that contained 1 mM EDTA.

Viral RNA was isolated from the purified preparation of virus by incubation with proteinase K at 1 mg/ml and SDS at 1% (w/v) at 37 °C for 20 min, with two subsequent extractions with a mixture of phenol and chloroform (1:1, v/v) and precipitation in ethanol. RNA was further purified by affinity chromatography on oligo(dT)-cellulose [13] and its purity was examined by electrophoresis on a 1% agarose gel. The RNA was used as the template for oligo(dT)-primed synthesis of cDNA, which was followed by second-strand synthesis [7], using a commercial kit (Amersham, U.K.). The double-stranded cDNA was ligated to *Sma*I-cut, dephosphorylated Bluescript II SK (+) (Stratagene). The DNA was used to transform competent *E. coli* JM109 cells. A recombinant plasmid (PL50) with a cDNA insert of approximately 1.4 kb was selected and used for Northern blot analysis with the ECL direct nucleic acid labeling and detection systems (Amersham) according to the manufacturer's instructions. The cDNA insert was excised by digestion with *Bam*HI and *Hin*dIII and ligated to *Bam*HI- and *Hin*dIII-cut Bluescript II KS (+) (Stratagene) to obtain a cDNA clone with the sequence in the opposite orientation. The two cDNA clones were digested with exonuclease III and mung bean nuclease (Takara, Kyoto, Japan) or S1 nuclease (Pharmacia) to produce two sets of nested deletion mutants for sequence analysis. The DNA sequences were determined by the dideoxynucleotide chain-termination method [19] on an automated DNA sequencer (model 377A; ABI). All parts of the cDNA were sequenced in both orientations. Sequence data were analyzed with the DNASIS system (version 7.0) from Hitachi Software Engineering Co. (Tokyo, Japan).

For analysis of the amino acid sequence, PLDMV CP was purified by SDS-PAGE [11], blotted onto a PVDF membrane (Japan Genetics, Tokyo, Japan) and was analyzed with an automated gas-phase protein sequencer (model 477A; ABI) as described by Kashiwazaki et al. [9].

Agarose gel electrophoresis of PLDMV RNA after purification by affinity chromatography on oligo (dT)-cellulose gave a single band of RNA of about 10 kb. A probe derived from clone PL50 hybridized specifically with this band of RNA (data not shown). A stretch of 32 adenosine residues was found at one end of PL50, suggesting that this clone contained cDNA that corresponded to the 3′-terminal region of the PLDMV RNA. The sequence of the 3′-terminal 1 404 nucleotides [excluding the poly(A) tail] of the PLDMV RNA, as determined from PL50, is shown in Fig. 1. The nucleotide sequence contains a long open reading frame (ORF) that potentially encodes a polypeptide of 397 amino acids, in frame 2 (Fig. 1). Other reading frames on this strand and on the complementary strands contain many stop codons and a few extended ORFs. The long ORF is open at the 5′ end and terminates with a UAG codon at positions 1 193–1 195, which is followed by a non-coding region of 209 nucleotides upstream of the 3′ poly(A) tail. A potential polyadenylation signal (UAUGU) [26] is present at positions 1 289–1 293.

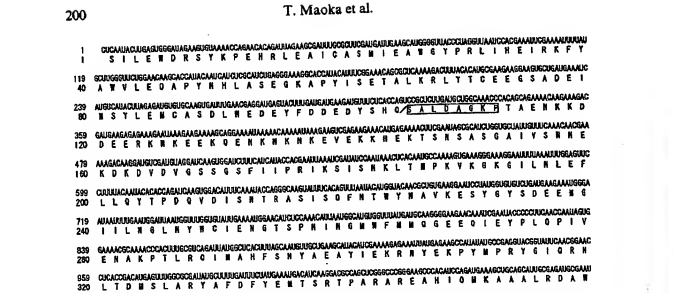

Fig. 1. Nucleotide sequence of the 3′-terminal 1404 nucleotides of PLDMV RNA. The predicted amino acid sequence encoded by the single long ORF is shown below the nucleotide sequence. A slash indicates a putative cleavage site between the NIb and CP proteins. The stop codon at the end of the long ORF is indicated by an asterisk. A potential polyadenylation signal (UAUGU) is underlined. The N-terminal amino acids determined by direct sequencing of CP are boxed. This sequence has been deposited with the EMBL/GenBank/DDBJ databases under accession no. D50082

Complete or partial genomic sequences have been reported for a number of potyviruses. The potyvirus genome encodes a single large polyprotein from which functional proteins are produced by cleavage at specific sites [20]. Comparison of the putative product encoded by the ORF of PLDMV with the C-terminal part of polyproteins of other potyviruses indicated that the putative product contains the whole presumptive CP at the C-terminus and a part of the NIb protein upstream of CP. A search for consensus cleavage sites (Q/S, Q/A and Q/G) reported for potyviruses revealed two candidates, Q/A at positions 45/46 and Q/S at positions 104/105, for the site of cleavage between the NIb and CP proteins of PLDMV. Direct sequencing of the PLDMV CP yielded its N-terminal sequence, namely, SALDAGKP, which is located in the putative product at positions 105–112. The CP protein contains the DAG motif (at positions 108–110), which has been reported to be located at the N-terminus of potyvirus CPs and is essential for transmissibility by aphid of the virus [20]. These results indicate that the Q/S sequence probably corresponds to the cleavage site between the NIb and CP proteins of PLDMV. In PLDMV, DAG motif is very close to the N-terminus of CP when compared with other potyviruses [20]. The PLDMV CP consists of 293 amino acids with a calculated M_r of 33 277. This value is lower than the M_r (37 000; data not shown) of CP

Table 1. The extent of amino acid sequence similarity (%) between capsid proteins (CPs) of PLDMV and other potyviruses

	PRSV-P (HA)	PRSV-P (MILD)	PRSV-W (USA)	PRSV-W (Aust)	PVY	SbMV	TEV	TVMV	ZYMV	WMV2
PLDMV	55	59	58	59	58	57	56	49	57	56
PRSV-P(HA)		92	91	91	57	53	56	51	55	53
PRSV-P(MILD)			97	97	60	56	59	53	57	56
PRSV-W(USA)				98	60	55	58	53	57	56
PRSV-W(Aust)					61	55	58	52	57	56
PVY						61	63	55	61	62
SbMV							60	53	71	84
TEV								59	63	59
TVMV									55	53
ZYMV										71
WMV2										

References for sequences are as follows: PRSV-P(HA) [25]; PRSV-P(MILD) [17]; PRSV-W(USA) [17]; PRSV-W(Aust) [2]; PVY [18]; SbMV [8]; TEV [1]; TVMV [4]; ZYMV [6]; WMV2 [5]. Percent similarity was calculated with the DNASIS program

estimated by SDS-PAGE of a purified preparation of PLDMV. In some other potyviruses, M_r values of CPs calculated from amino acid sequences are also smaller than the M_r values estimated by SDS-PAGE [20].

The CPs of members of the genus *Potyvirus* have a bimodal distribution of amino acid sequence similarity. There is 90% to 99% similarity between strains, and 38% to 71% similarity between distinct viruses [20]. The deduced CP of PLDMV shares 55% to 59% sequence similarity with the CPs of four isolates of PRSV, among which amino acid sequence similarity was found to be 91% to 98% (Table 1). The low degree of similarity at the amino acid level between the CPs of PLDMV and PRSV is consistent with the absence of an obvious serological relationship between them [10, 12]. The 3′ non-coding region (3′NCR) of PLDMV RNA (1 196–1 404) also showed low similarity (ca. 30%) to the 3′NCR of RNAs of four isolates of PRSV (data not shown). These differences among sequence support the hypothesis that PLDMV and PRSV are different viruses [20].

CP amino acid sequence similarity between PLDMV and six other potyviruses varied from 49% to 58% (Table 1), and that between PLDMV and 40 isolates of 33 distinct aphid-transmitted potyviruses [14, 16, 20] also varied from 49% to 59% (data not shown). The results are similar to those reported for distinct potyviruses [20]. These results indicate that PLDMV is a distinct virus in the genus *Potyvirus*.

Since a resistance gene to PRSV-P has not been identified in papaya, a cross-protection mechanism has been used to control devastating infections by PRSV-P [22, 23]. However, cross-protection with attenuated strains of PRSV-P has proved only partial or ineffective in some areas [24]. The failure or ineffectiveness of cross-protection has been attributed to variability of PRSV. Although the distribution of PLDMV is unknown, the presence of PLDMV may also be involved in decreasing the effectiveness of the cross-protection mechanisms in some locations. In Okinawa, PLDMV occurs more frequently than PRSV-P [12]. It is important to prove the absence of PLDMV in areas affected by papaya ringspot. Otherwise, evaluation of potential control measures may be inadequate.

References

1. Allison R, Johnston RE, Dougherty WG (1986) The nucleotide sequences of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polyprotein. Virology 154: 9–20
2. Bateson M, Dale J (1992) The nucleotide sequence of the coat protein gene and 3′ untranslated region of papaya ringspot virus type W (Aust). Arch Virol 123: 101–109
3. Bateson MF, Henderson J, Chaleeprom W, Gibbs AJ, Dale JL (1994) Papaya ringspot potyvirus: isolate variability and the origin of PRSV type P (Australia). J Gen Virol 75: 3547–3553
4. Domier LL, Franklin KM, Shahabuddin M, Hellmann GM, Overmeyer JH, Hiremath mST, Siaw MFE, Lomonossoff GP, Shaw JG, Rhoads RE (1986) The nucleotide sequence of tobacco vein mottling virus RNA. Nucleic Acids Res 14: 5417–5430

5. Frenkel MJ, Ward CW, Shukla DD (1989) The use of 3' non-coding nucleotide sequences in the taxonomy of potyviruses: application to watermelon mosaic virus 2 and soybean mosaic virus-N. J Gen Virol 70: 2775–2783
6. Gal-On A, Antignus Y, Rosner A, Raccah B (1990) Nucleotide sequence of the zucchini yellow mosaic virus capsid-encoding gene and its expression in *Escherichia coli*. Gene 87: 273–277
7. Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. Gene 25: 263–269
8. Jayaram C, Hill JH, Miller WA (1992) Complete nucleotide sequences of two soybean mosaic virus strains differentiated by response of soybean containing the *Rsv* resistance gene. J Gen Virol 73: 2067–2077
9. Kashiwazaki S, Nomura K, Kuroda H, Ito K, Hibino H (1992) Sequence analysis of the 3'-terminal halves of RNA 1 of two strains of barley mild mosaic virus. J Gen Virol 73: 2173–2181
10. Kawano S, Yonaha T (1992) The occurrence of papaya leaf-distortion mosaic virus in Okinawa. Tech Bull of FFTC 132: 13–23. Food and Fertilizer Technology Center for the Asian and Pacific Region, Taipei, Taiwan, Republic of China
11. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227: 680–685
12. Maoka T, Kawano S, Usugi T (1995) Occurrence of the P strain of papaya ringspot virus in Japan. Ann Phytopathol Soc Japan 61: 91–94
13. Nakazato H, Edmonds M (1974) Purification of messenger RNA and heterogeneous nuclear RNA containing poly(A) sequences. Methods Enzymol 29: 431–443
14. Pappu SS, Pappu HR, Rybicki EP, Niblett CL (1994) Unusual amino-terminal sequence repeat characterizes the capsid protein of dasheen mosaic potyvirus. J Gen Virol 75: 239–242
15. Purcifull DE, Edwardson JR, Hiebert E, Gonsalves D (1984) Papaya ringspot virus. CMI/AAB Descriptions of Plant Viruses, no. 292
16. Puurand U, Makinen K, Paulin L, Saarma M (1994) The nucleotide sequence of potato virus A genomic RNA and its sequence similarities with other potyviruses. J Gen Virol 75: 457–461
17. Quemada H, L'Hostis B, Gonsalves D, Reardon IM, Heinrikson R, Hiebert EL, Sieu LC, Slightom JL (1990) The nucleotide sequences of the 3'-terminal regions of papaya ringspot virus strains W and P. J Gen Virol 71: 203–210
18. Robaglia C, Durand-Tardif M, Tronchet M, Boudazin G, Astier-Manifacier S, Casse-Delbart F (1989) Nucleotide sequence of potato virus Y (N strain) genomic RNA. J Gen Virol 70: 935–947
19. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467
20. Shukla DD, Ward CW, Brunt AA (1994) The potyviridae. CAB International, Wallingford, Oxon
21. Wang CH, Bau HJ, Yeh SD (1994) Comparison of the nuclear inclusion b protein and coat protein gene of five papaya ringspot virus strains distinct in geographic origin and pathogenicity. Phytopathology 84: 1205–1210
22. Wang HL, Yeh SD, Chiu RJ, Gonsalves D (1987) Effectiveness of cross-protection by mild mutants of papaya ringspot virus for control of ringspot disease of papaya in Taiwan. Plant Dis 71: 491–497
23. Yeh SD, Gonsalves D (1984) Evaluation of induced mutants of papaya ringspot virus for control by cross protection. Phytopathology 74: 1086–1091

24. Yeh SD, Gonsalves D, Wang HL, Namba R, Chiu RJ (1988) Control of papaya ringspot virus by cross protection. Plant Dis 72: 375–380
25. Yeh SD, Jan FJ, Chiang CH, Doong TJ, Chen MC, Chung PH, Bau HJ (1992) Complete nucleotide sequence and genetic organization of papaya ringspot virus RNA. J Gen Virol 73: 2531–2541
26. Zaret KS, Sherman F (1982) DNA sequence required for efficient transcription termination in yeast. Cell 28: 563–573

Authors' address: Dr. T. Maoka, Okinawa Subtropical Station of the Japan International Research Center for Agricultural Sciences, Maezato, Ishigaki, Okinawa 907, Japan.

Received May 16, 1995